A Survey of GPT-3 Family Large Language Models Including ChatGPT and GPT-4

Year: 2023 | *Citations: 324* | *Authors: Katikapalli Subramanyam Kalyan*

Abstract

Large language models (LLMs) are a special class of pretrained language models obtained by scaling model size, pretraining corpus and computation. LLMs, because of their large size and pretraining on large volumes of text data, exhibit special abilities which allow them to achieve remarkable performances without any task-specific training in many of the natural language processing tasks. The era of LLMs started with OpenAI GPT-3 model, and the popularity of LLMs is increasing exponentially after the introduction of models like ChatGPT and GPT4. We refer to GPT-3 and its successor OpenAI models, including ChatGPT and GPT4, as GPT-3 family large language models (GLLMs). With the ever-rising popularity of GLLMs, especially in the research community, there is a strong need for a comprehensive survey which summarizes the recent research progress in multiple dimensions and can guide the research community with insightful future research directions. We start the survey paper with foundation concepts like transformers, transfer learning, self-supervised learning, pretrained language models and large language models. We then present a brief overview of GLLMs and discuss the performances of GLLMs in various downstream tasks, specific domains and multiple languages. We also discuss the data labelling and data augmentation abilities of GLLMs, the robustness of GLLMs, the effectiveness of GLLMs as evaluators, and finally, conclude with multiple insightful future research directions. To summarize, this comprehensive survey paper will serve as a good resource for both academic and industry people to stay updated with the latest research related to GPT-3 family large language models.